Supporting Information

Litovchick and Szostak 10.1073/pnas.0805837105

SI Text

Selection Library and Cyclic Peptide-mRNA Fusion Synthesis. The initial DNA library was synthesized by the Keck facility at Yale University (Fig. S1). The library encodes a T7 promoter, TMV enhancer, start codon ATG, X-codon RRR (which could translate into Arg, Lys, Glu, Asp, or Gly). It also includes a random 30-nt region $(NNB)_{10}$, encoding a random 10-mer peptide, flanked by cysteine codons TGC. The choice of NNB random triplets decreases the number of stop codons in the random portion. The 3′ constant region encodes a GSVG spacer and 6 histidines (his_6-tag). The nine nucleotides downstream of the his_6-tag are complementary to the 2′-O-Me RNA portion of the cross-linking puromycin-terminated oligonucleotide. They encode the tripeptide HRL, which is followed by a TAG stop codon. We have found that the mRNA is translated up to the stop codon, so all peptide sequences contained the HRL tripeptide at their C termini. The 3′ constant region contains two additional out-of-frame stop codons.

The double-stranded DNA library was prepared by primer extension by using the 3LIBHIS primer and purified by 10% native PAGE. The DNA library was *in vitro* transcribed by T7 RNA polymerase to generate the mRNA library. The puromycin oligo (25 μM) was annealed to the mRNA library (10 μM) in 2 ml of buffer, containing 20 mM Hepes, pH 7.4, 100 mM KCl, 1 mM spermidine, and 1 mM EDTA, and was cross-linked by UV 365-nm irradiation as described (2). Photo-cross-linked mRNA was purified by 8% denaturing 8 M urea PAGE. For round 1 *in vitro* translation and fusion formation was performed in 4 ml of wheat germ (WG) extract (Promega) with 0.5 μM cross-linked mRNA for 1 h at 30°C. The translation reaction was then stopped by dilution with 940 μl of 2.5 M KCl and 260 μl of 1 M $MgCl_2$, incubated for 15 min at room temperature, and then frozen at −20°C to facilitate fusion formation. The sample (5.2 ml), containing an estimated 0.2 nmol of fusions, was then diluted to 20 ml with buffer A [1 M NaCl, 20 mM Tris·HCl, pH 8.0, 0.2% Tween-20, 0.2 mM Tris(carboxyethyl)-phosphine (TCEP), 20 mM EDTA] and applied to 500 mg of oligo(dT) cellulose (NEB), soaked in 15 ml of buffer A for 30 min at 4°C with gentle shaking. The resin with the absorbed fusions was loaded on a 15-ml disposable plastic column (Bio-Rad) and washed extensively with buffer A. The column was equilibrated with the cyclization buffer (660 mM KCl, 20 mM Tris, pH 8.0, 0.2 mM TCEP, and 3 mM dibromo-*m*-xylene in 30% acetonitrile/70% water mixture) and was incubated for 1 h with gentle shaking. After cyclization, the column was washed with 3 column volumes of 300 mM KCl, 20 mM Tris·HCl, pH 7.2, and eluted with 2 mM Tris, pH 7.2, in 500-μl fractions. The cyclized fusion were collected, concentrated, and then dissolved in 6 M guanidine HCl/50 mM sodium phosphate buffer, pH 8.0, and applied to a Ni-NTA column (Qiagen). The column was washed with the same buffer, then with 50 mM sodium phosphate, pH 8.0, 20 mM imidazole buffer, and finally eluted with 400 mM imidazole. The fusions were ethanol-precipitated and reverse-transcribed by using RT primer (Fig. S1*B*).

In all rounds of selection, PCR amplification of the eluate was performed with 5T7TMVLib and 3LIBHIS primers, as indicated in Fig. S1.

***In Vitro* Translation and Cyclization of Peptide 6B4.** *In vitro* translation of the 6B4 peptide was performed in a reconstituted *E. coli* translation mixture (PURE system; 3) in the presence of ^{35}S methionine, with cyclization on a Ni-NTA column (4). Translation buffer contained: 10 mM Tris·HCl, 10 mM $Mg(OAc)_2$, 100 mM NH_4Cl, pH 7.5 (at 37°C), and reactions (50 μl) were typically incubated for 1 h at 37°C. The reactions were diluted with 100 μl of wash buffer (50 mM Tris·HCl, 300 mM NaCl, pH 8) and 50 μl of NTA-agarose beads (Qiagen) were added. After incubation (30 min) the agarose beads were washed (wash buffer + 0.2 mM TCEP) then treated with 5 mM dibromo-*m*-xylene (Aldrich) and 0.2 mM TCEP in a 1:3 acetonitrile/50 mM Tris·HCl buffer, pH 8.0, for 1 h at room temperature. The peptides were eluted with 0.2% TFA and the yield was quantified by liquid scintillation counting of ^{35}S methionine. Typical yields were 10–15 pmol from a 50-μl reaction. For MALDI-TOF analysis, peptides were desalted and concentrated by reverse-phase microchromatography (C18 Zip Tips, Millipore) and eluted with a 55% acetonitrile, 0.1% TFA solution saturated with α-cyano-4-hydroxycinnamic acid. We used oxidized insulin chain B (M_r 3,495) as a mass standard.

Synthetic 6B4 Peptide. The 27-residue peptide MKCSRGIRCAGVLCGSVGHHHHHHHRL (6B4), the 8-residue variant of 6B4 referred to as 6B48, KCSRGIRC, and the 27-mer 6B4 labeled at its N terminus with the 6-isomer of fluorescein isothiocyanate (Fl-6B4) were synthesized by using F-moc chemistry and purified by GenScript Corp. Full-length 6B4 was cyclized with dibromo-*m*-xylene on a Ni-NTA column as described above, producing a bicyclic derivative of the peptide, 6B4C. Peptide 6B48 was cyclized by reaction with 1.1 equiv of dibromo-*m*-xylene in 1:3 acetonitrile/50 mM Tris·HCl buffer, pH 8.0, 0.2 mM TCEP for 1 h at room temperature, producing monocyclic 6B48C. The peptides were purified on a 250 × 4.6 mm reverse-phase C-18 HPLC column in a gradient of 10–50% acetonitrile with 0.1% TFA (6B4C and 6B48C) or a 10–70% acetonitrile gradient in 0.1 M TEAB, pH 7.8 (Fl-6B4). The molecular weight of the peptides were determined by MALDI-TOF MS.

Equilibrium Ultrafiltration Measurements of the Dissociation Constants by Using Fluorescently Labeled 6B4 Peptide. Fluorescent peptide Fl-6B4 was used as a probe for solution binding and competition experiments. A sample of 200 μl of 0.4–2 nM Fl-6B4 in a buffer (20 mM KHEPES, pH 7.4, 300 mM NaCl, 5 mM $MgCl_2$, 2 mM $CaCl_2$, 0.025% Triton X-100, and 0.5% DMSO), was incubated for 1 h with an increasing concentration of freshly transcribed and purified HCV IRES RNA. Equilibrated solutions (200 μl) were transferred to YM-30 spin filters (Millipore) and centrifuged for 15 s at 13,500 × *g* to yield about 10 μl of filtrate, which was discarded. After an additional spin of 1.5 min, top and bottom samples (about 90 μl each) were transferred into a 80-μl quartz cuvette, and fluorescence spectra were collected in a range of 505–550 nm (slit width, 10 nm) on a Cary Eclipse spectrofluorimeter (Varian) at excitation wavelength 495 nm (slit width, 5 nm) at high PMT voltage (800 V); smoothing, using the moving average method with factor 6. The spectrum of buffer without fluorescent tracer was collected each time and subtracted from a fluorescence spectrum. The average value was calculated for the emission range of 517–522 nm and was used to quantify the amount of fluorescent peptide in each sample. Seven to fifteen values for the fraction of bound ligand were measured and plotted against IRES concentration. The solution binding data then were fit to the following binding equation by using KaleidaGraph (5): $Y = B + F * I/(I + K)$, where Y = the fraction of labeled peptide bound to IRES RNA, I = the HCV

IRES concentration, B = nonspecific binding of the peptide to filter, F = maximum fraction of counts that can be bound, K = the dissociation constant (m1, F and K are parameters fitted through nonlinear regression).

For competitive binding experiments, to 200 μl of 0.4–0.625 nM Fl-6B4 and 15–18 nM IRES (determined to give $\approx$70% binding), preequilibrated for 1 h in the same buffer, increasing concentrations of competing peptides were added and incubated for an additional 1 h. The samples were then treated as described above. The solution binding data using unlabeled competitor peptides were fit to the following binding equation (5): $Y = \{K/[IF - 0.5*\{(IF + C + X) - ((IF + C + X)^2 - 4IFC)^{1/2}\}] + 1\}^{-1}$, where Y = the fraction of Fl-6B4 bound to IRES (normalized to 1 at the absence of competitor peptides), K = the determined dissociation constant for Fl-6B4, X = the dissociation constant for the competitor peptide, I = the IRES concentration, C = the unlabeled competitor concentration, and F = the fraction of the IRES that is correctly folded.

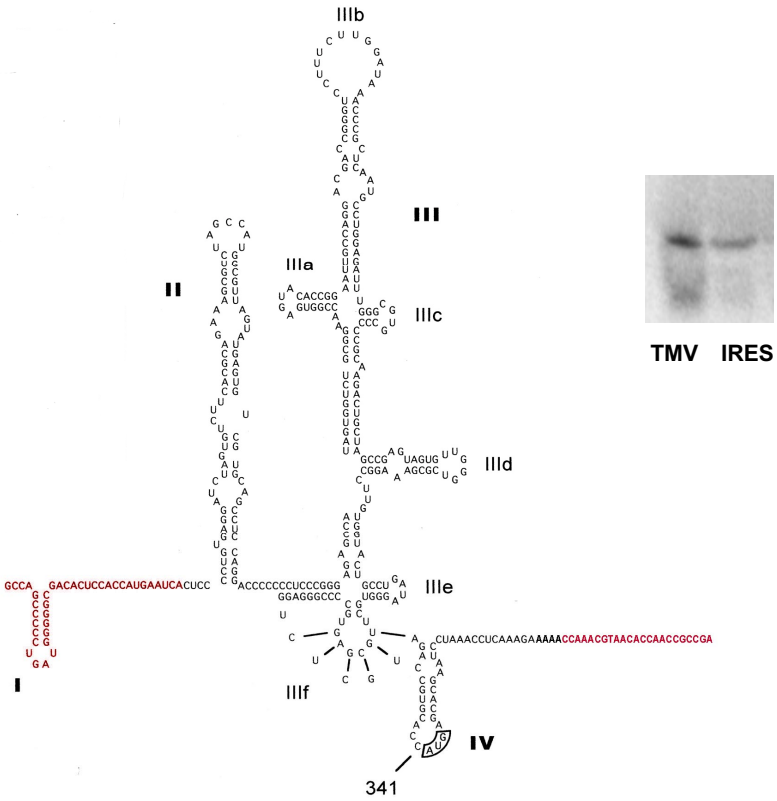
HIV RRE RNA and Class I Ligase Ribozyme. HIV Rev responsive element (RRE, 247-mer) template was PCR-amplified from the pNL4–3 plasmid (National Institutes of Health AIDS Research & Reference Reagent program) that contains a full HIV provirus copy, using primers T7RRE5, GCTAATACGACTCACTATAGAGCAGTGGGAATAGGAGC (underlined portion is T7 promoter) and 3RRE, AGGAGCTGTTGATCCTTTAGGTATC. RRE RNA was *in vitro* transcribed by using T7 RNA polymerase from the PCR template.

Purified 140-nt-long class I ligase ribozyme RNA (6) was a generous gift from David M. Shechner and Dr. David P. Bartel (Whitehead Institute, Massachusetts Institute of Technology).

Affinities of the peptides to noncognate RNA targets such as HIV RRE RNA, class I ligase ribozyme, as well as rRNA, tRNA, and CW mRNA, were measured as described above.

1. Honda M, Beard MR, Ping LH, Lemon SM (1999) A phylogenetically conserved stem-loop structure at the 59 border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J Virol* 73:1165–1174.
2. Liu R, Barrick JE, Szostak JW, Roberts RW (2000) Optimized synthesis of RNA-protein fusions for in vitro protein selection. *Methods Enzymol* 318:268–293.
3. Shimizu Y, *et al.* (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol* 19:751–755.
4. Seebeck FP, Szostak JW (2006) Ribosomal synthesis of dehydroalanine-containing peptides. *J Am Chem Soc* 128:7150–7151.
5. Sazani PL, Larralde R, Szostak JW (2004) A small aptamer with strong and specific recognition of the triphosphate of ATP. *J Am Chem Soc* 126:8370–8371.
6. Bartel DP, Szostak JW (1993) Isolation of new ribozymes from a large pool of random sequences. *Science* 261:1411–1418.
7. Coulon A, *et al.* (2002) Modes of membrane interaction of a natural cysteine-rich peptide: Viscotoxin A3. *Biochim Biophys Acta* 1559:145–159.

A



Peptide library

B

MXCNNNNNNNNNNCGSVGHHHHHHHHRL

DNA library

T7 promoter **TMV enhancer**
GCTAATACGACTCACTATA GGACAATTACTATTTACAATTACA

fMet X C N N N N N N N N N **C**
atg RRR **TGC** NNB NNB NNB NNB NNB NNB NNB NNB NNB NNB **TGC**

G S V G H H H H H H H R L Stop
GG***T AG***C G***TA G***GT CACCATCACCACCATCAC CAC CGG CTA**TAG**
(out of frame Stop codons)

Primers:

5T7TMVLIB: GCTAATACGACTCACTATAGGACAATTACTATTTACAATTACAATG
3LIBHIS: CTATAGCCGGTGGTGATGGTGGTGATGGTGACCTACGCTACCGCA
RT: TTTTTTTTTTTTTTTGTGATGGTGGTGATGGTG

Crosslinking puromycin oligo:

X**(tagccggtg)** AAA AAA AAA AAA AAA ZZ ACC P

Fig. S1. Selection target and library design. (*A*) Secondary structure of HCV IRES RNA, genotype 1b, nucleotides 1–395 (adapted from ref. 1) used as the selection target. Portions highlighted in red were added to the original clone 40–372 by PCR. The 5′-terminal 40 nucleotides were added to facilitate cotranscriptional folding of IRES. The 3′-terminal 23 nucleotides were added to provide higher stability and to increase the activity of the IRES as well as for the translation of an N-terminal 18-mer peptide of the HCV core protein used for preliminary tests of IRES function. (*Inset*) Translation of the IRES construct in Rabbit Reticulocyte Lysate. For the control the same sequence encoding for first 18 aa of the HCV core protein was cloned under the control of the TMV enhancer. (*B*) Design of the library and primers used for the selection. For the cross-linking puromycin-terminated oligonucleotide: X = psoralen C6, (lowercase) = 2′OMe ribonucleotides, Z = spacer 9, P = puromycin, stretch of As and ACC are deoxyribonucleotides.

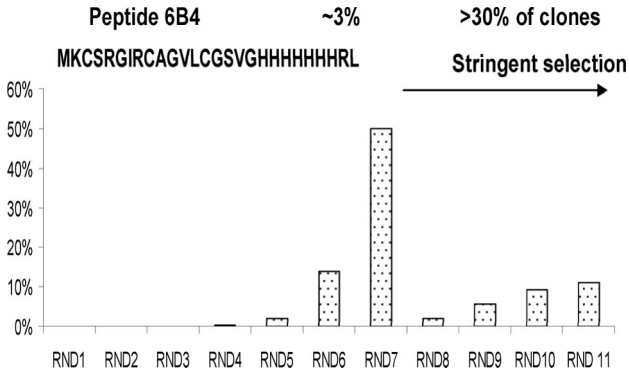


Fig. S2. Progress of the selection: fraction of ^{35}S-radiolabeled input library specifically eluted from the selection column in each round. Selection stringency was increased after round 8.

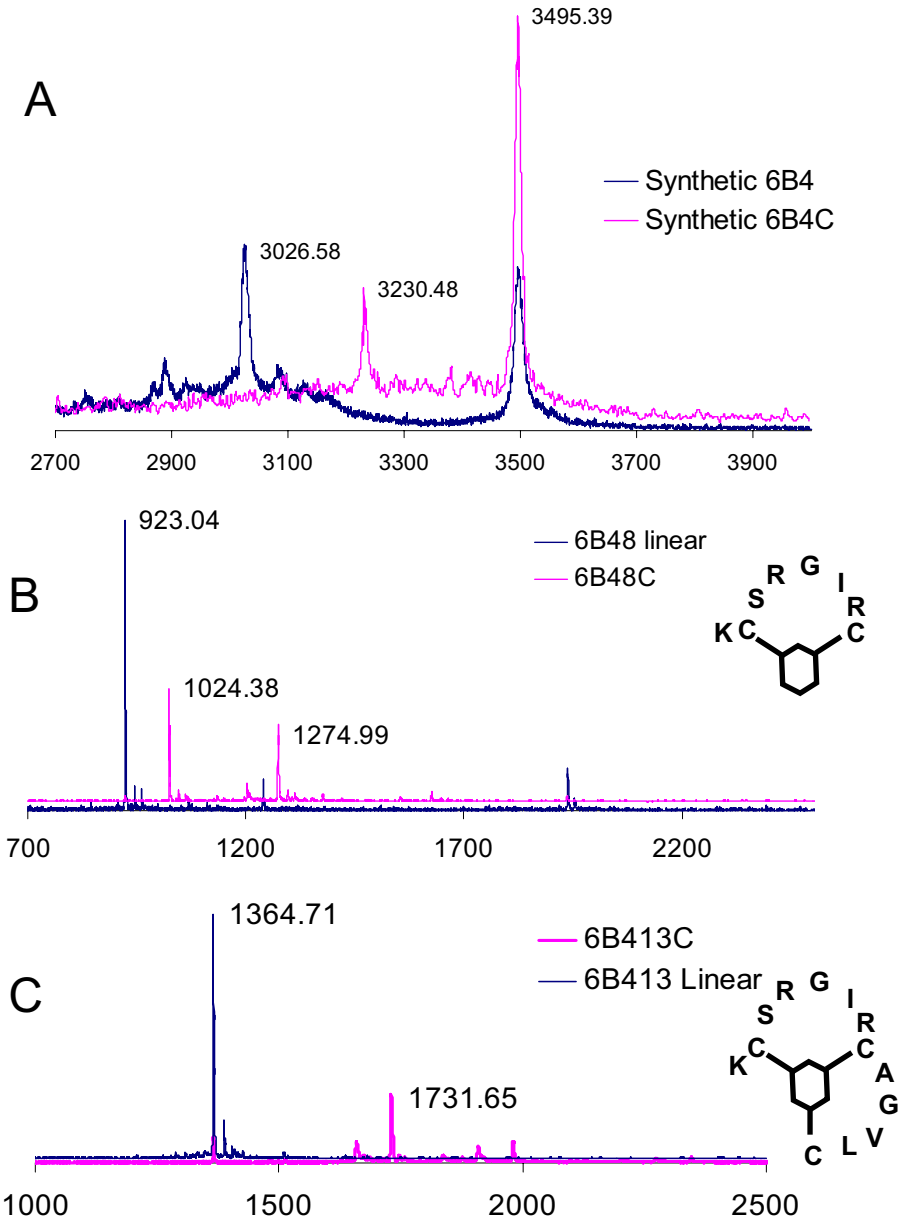


Fig. S3. MALDI-TOF spectra of several peptides used in the study. (*A*) Synthetic (GenScript) 6B4 peptide. Because the peptide lacks N-terminal formylation, the mass is 28 units lower than that of *in vitro* translated 6B4. Two xylene moieties are added by reaction with dibromo-*m*-xylene. Blue trace: linear peptide (observed mass, 3,026.58; expected mass, 3,026.6). Red trace: cyclic peptide (observed mass, 3,230.48; expected mass, 3,230.6). The peak at mass 3,495.39 is an internal standard (insulin chain B). (*B*) MALDI-TOF spectra of 6B48 (KCSRGIRC) in linear and cyclic form. The peptide was cyclized in solution by reaction with dibromo-*m*-xylene. The cyclic product (observed mass, 1,024.38; expected mass, 1,024.13) and a TCEP adduct (observed mass,1,274.99; expected mass, 1,274.2) are observed. (*C*) MALDI-TOF spectra of 6B413 (KCSRGIRCAGVLC) in linear and cyclic forms. The peptide represents the random portion of the selected 6B4 peptide. The cyclization of 6B413 was performed with tribromo mesitylene (Aldrich) to engage all three cysteine moieties without producing a mixture of products, under conditions similar to the dibromo-*m*-xylene reaction. Blue trace: linear 6B413 peptide (observed mass, 1,364.7; expected, 1,365.72); pink trace: bicyclic 6B413C (observed mass, 1,731.65; expected, 1,482.7 + 250 = 1,732.72 as TCEP adduct). All solution cyclization reactions produced mass spectra containing a TCEP adduct (+250), whereas cyclization of the peptide immobilized on a Ni- NTA column did not show a TCEP adduct on the mass spectrum.

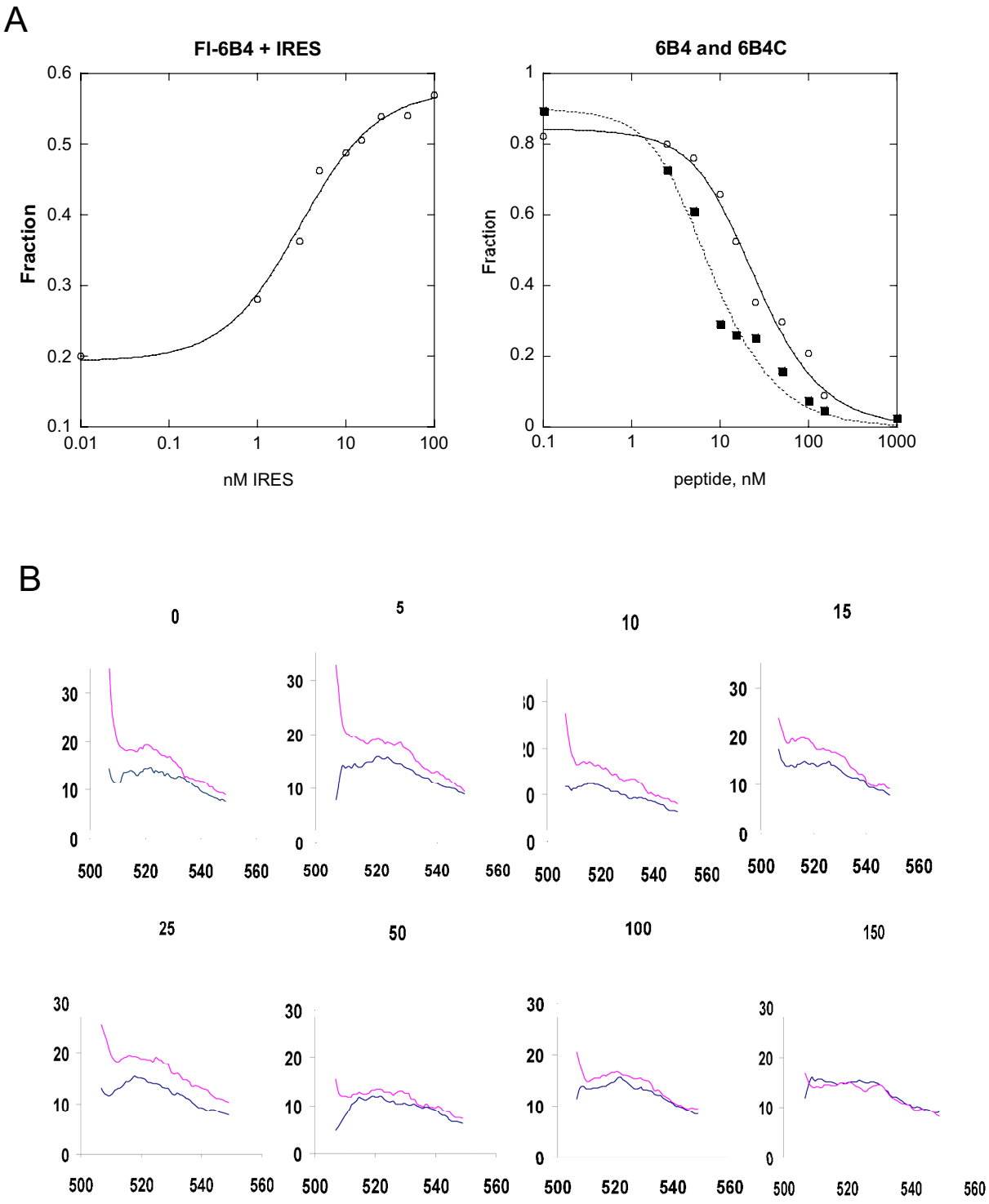


Fig. S4. Fluorescent equilibrium ultrafiltration assay. (*A*) Dissociation constants measured by equilibrium ultrafiltration using Fl-6B4. (*Right*) Binding of HCV IRES RNA to Fl-6B4 (2 nM). (*Left*) Competitive binding of 6B4 (open circles) and 6B4C (filled squares) to IRES RNA (15 nM) by using Fl-6B4 (2 nM) as a tracer. (*B*) Examples of spectra collected in equilibrium ultrafiltration assay by using Fl-6B4 as a tracer: Competition of 8-mer cyclic 6B48C with Fl-6B4 (0.4 nM) + IRES (15 nM). Concentration of the competitor peptide is indicated in titles of the graphs (nM).

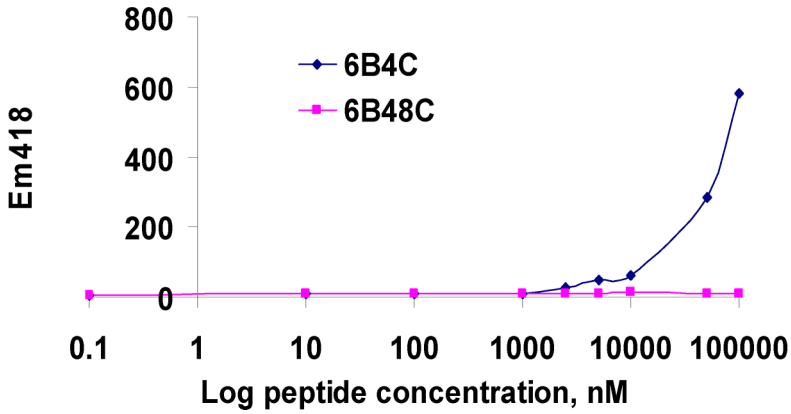


Fig. S5. Peptide aggregation measured by *N*-phenyl 1-naphthylamine fluorescence (7). A stock solution of *N*-phenyl 1-naphthylamine (NPN, Aldrich) in ethanol was diluted with a buffer (20 mM KHEPES, pH 7.4, 300 mM NaCl, 5 mM $MgCl_2$, 2 mM $CaCl_2$) to a final concentration of 50 nM. Different concentrations of peptides were added to the solutions and incubated at room temperature for 45 min. Spectra were recorded on a Cary spectrofluorimeter in the range of 380–550 nm (emission slit width, 10 nm) at excitation wavelength of 340 nm (slit width, 5 nm) filtered at 360–1,100 nm, PMT voltage high (800 V), smoothing by running-average method, factor 6. (Blue line) A change of fluorescence behavior of NPN in the presence of the cyclic 27-mer 6B4C indicated aggregation of the peptide. (Pink line) At the same range of concentrations of cyclic 8-mer 6B48C no changes in NPN fluorescence occurs, indicating a lack of aggregation of 6B48C.

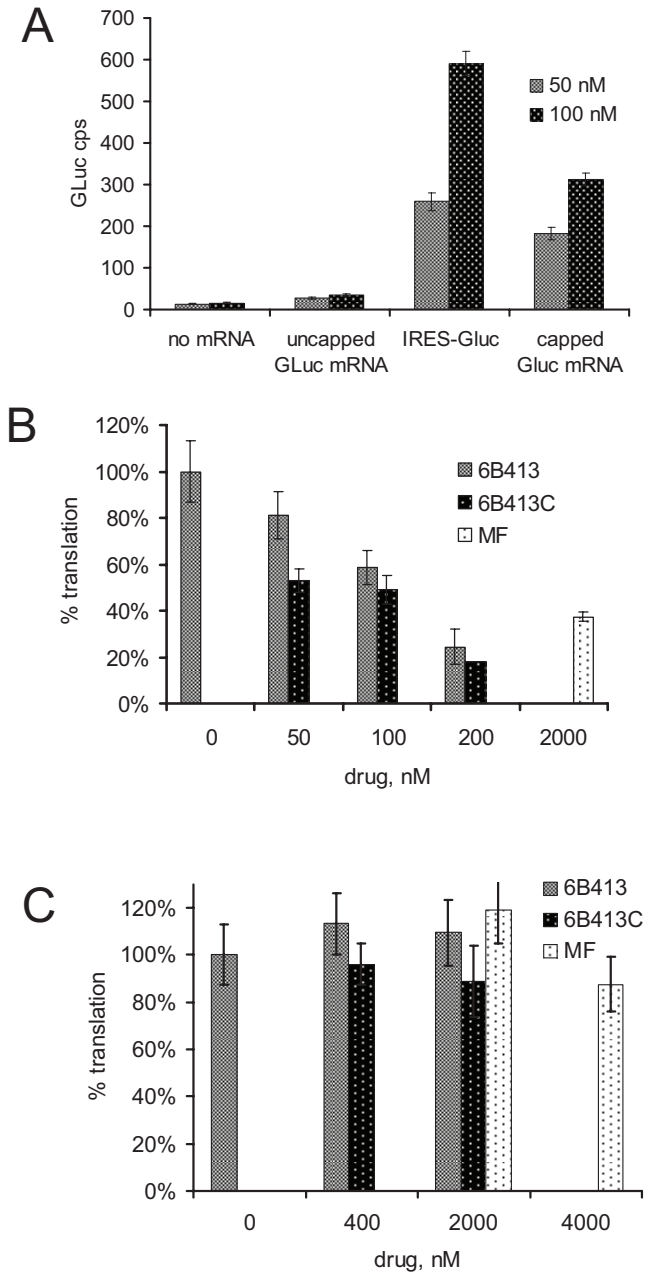


Fig. S6. *In vitro* translation experiments with IRES-Gluc construct. (*A*) Translation of IRES-GLuc and uncapped or capped GLuc mRNA in HeLa S10 extract. Uncapped GLuc mRNA translates poorly, whereas capped GLuc and IRES-Gluc construct translate well. The slightly lower translation yield of capped GLuc mRNA compared with IRES-GLuc may reflect the 50–70% efficiency of cap analog incorporation during *in vitro* transcription of the construct by T7 RNA polymerase. (*B*) Translation of 50 nM IRES-GLuc mRNA in HeLa S10 extract in the presence of a linear 13-mer 6B413 or cyclic 6B413C peptides KCSRGIRCAGVLC and mifepristone (MF). (*C*) Translation of 100 nM capped leader-GLuc in HeLa extract in the presence of different concentrations of 6B413 peptide in linear and cyclic form and mifepristone (MF).

KCSRGIRC

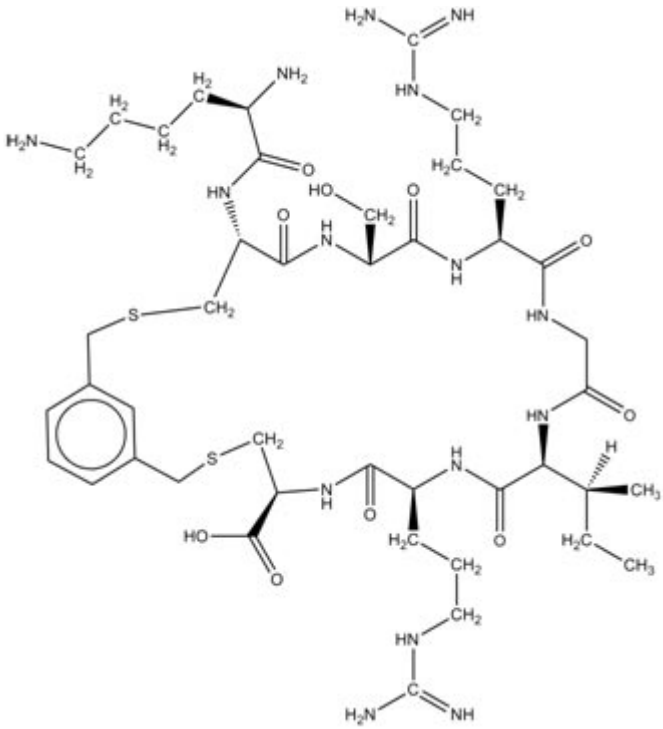


Fig. S7. Structure of the 8-mer cyclic peptide 6B48C.

Table S1. Conditions for different rounds of the selection

Round	NaCl Binding/elution	Arginine, mM	Competitor RNA	Elution of the fusions
1–2	0.75 M	10	TYR 20 μM	8 M urea and 50 mM NaOH
3–7	0.5 M/200 mM	15	TYR 40–80 μM first pre-elution with 100 μM TYR	Second elution with 10–12 μM soluble IRES for 1–2 h
8–11	0.5 M/200 mM	None	TYR 80–100 μM first pre-elution with rRNA, tetrahymena intron, 319 mRNA, 2 h (total RNA concentration up to 30 μg/μl)	Second pre-elution with 10–12 μM soluble IRES 2 h; third elution with 5–10 μM soluble IRES 12 h

Other Supporting Information

Appendix (PDF)